



GFIT2: an experimental algorithm for vertical profile retrieval from near-IR spectra

Brian J. Connor¹, Vanessa Sherlock², Geoff Toon³, Debra Wunch⁴, and Paul O. Wennberg⁴

¹BC Consulting Limited, Martinborough, New Zealand

²National Institute for Water and Atmospheric Research, Wellington, New Zealand

³Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

⁴California Institute of Technology, Pasadena, CA, USA

Correspondence to: Brian J. Connor (bcconsulting@xtra.co.nz)

Received: 27 August 2015 – Published in Atmos. Meas. Tech. Discuss.: 23 November 2015

Revised: 29 June 2016 – Accepted: 4 July 2016 – Published: 2 August 2016

Abstract. An algorithm for retrieval of vertical profiles from ground-based spectra in the near IR is described and tested. Known as GFIT2, the algorithm is primarily intended for CO₂, and is used exclusively for CO₂ in this paper. Retrieval of CO₂ vertical profiles from ground-based spectra is theoretically possible, would be very beneficial for carbon cycle studies and the validation of satellite measurements, and has been the focus of much research in recent years. GFIT2 is tested by application both to synthetic spectra and to measurements at two Total Carbon Column Observing Network (TCCON) sites. We demonstrate that there are approximately 3° of freedom for the CO₂ profile, and the algorithm performs as expected on synthetic spectra. We show that the accuracy of retrievals of CO₂ from measurements in the 1.61 μ (6220 cm⁻¹) spectral band is limited by small uncertainties in calculation of the atmospheric spectrum. We investigate several techniques to minimize the effect of these uncertainties in calculation of the spectrum. These techniques are somewhat effective but to date have not been demonstrated to produce CO₂ profile retrievals with sufficient precision for applications to carbon dynamics. We finish by discussing ongoing research which may allow CO₂ profile retrievals with sufficient accuracy to significantly improve the scientific value of the measurements from that achieved with column retrievals.

1 Motivation

Since 2004 the Total Carbon Column Observing Network (TCCON) has measured ground-based near-IR solar spectra. Their high spectral resolution and signal-to-noise ratio (SNR) allow high-precision measurements of total overhead column abundance of CO₂ and other gases, which provide constraints on the global carbon budget and also help validate satellite measurements using the same spectral bands. The standard analysis consists of least-squares spectral fitting to derive a multiplicative scale factor applied to an assumed (“a priori”) CO₂ profile shape. That analysis (“profile scaling”) provides column densities, from which dry-air average mole fractions (“XCO₂”) can be derived.

Allowing the a priori CO₂ profile shape to vary in the retrieval process (“profile retrieval”) has several potential advantages. For one, it can be shown that such an algorithm has more uniform sensitivity to CO₂ as a function of altitude and so should be less sensitive to bias from the a priori profile (Fig. 1). Secondly, it has more freedom to fit the observed spectrum, and thus generally will leave smaller residuals, and may help in understanding their origin. Finally, it would theoretically allow separation of the boundary layer from the rest of the column, helping to distinguish sources and sinks at continental and sub-continental scales (Fig. 2). Profile retrieval at the accuracy required is challenging, however. Profile retrieval requires that the factors that affect the spectral line shape (e.g., instrument line shape (ILS), spectroscopic widths) be accurately known. And since profile retrieval attempts to extract more information from the spectrum than

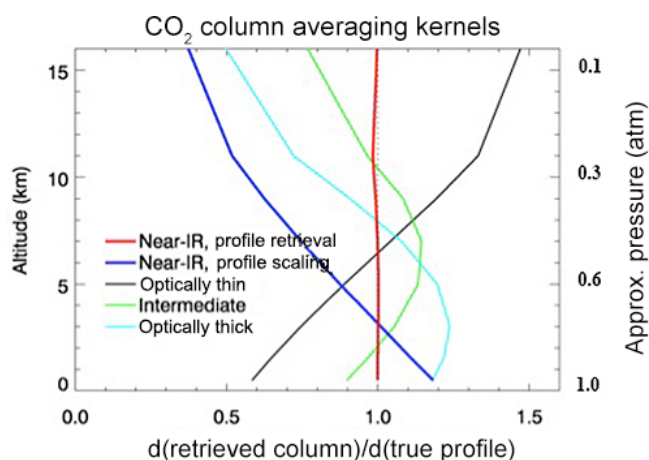


Figure 1. Column averaging kernels for simulated retrievals. “Profile retrieval” and “profile scaling” refer to a typical retrieval using GFIT2 on the 6220 cm^{-1} CO_2 band observed operationally. “Optically thin”, “optically thick”, and “intermediate” refer to an idealized, single, isolated spectral line. For these, we calculate a spectrum from a reference profile, perturb the profile at a single altitude, and calculate a new spectrum. We then perform a least-squares fit to this synthetic spectrum by deriving a scale factor for the reference profile. This scale factor, divided by the actual perturbation, produces a single element of the column averaging kernels shown.

profile scaling, there typically need to be a priori constraints to keep the retrieval stable (Rodgers, 1976; Solomon et al., 2000; Dohe, 2013).

It is worth noting that profile retrieval is unlikely to reduce uncertainties in determination of $X\text{CO}_2$. Wunch et al. (2015) showed that error due to the CO_2 a priori profile shape was less than $\sim 0.01\%$ at solar zenith angle (SZA) $< 70^\circ$, and increased only slowly at higher SZA, unless deliberately crude a priori profile shapes were chosen.

Recent work on profile retrieval development includes Kuai et al. (2013), who retrieved CO_2 in three tropospheric layers using the 1.6μ spectral band, and Dohe (2013), who studied a complete profile retrieval from measurements of the 2.1μ band.

This paper describes the experimental implementation and early tests of a profile retrieval algorithm for TCCON spectra. Its layout is as follows. In Sect. 2 we describe the standard algorithm used for TCCON and briefly discuss the history of profile retrieval and the chosen algorithm for similar measurements. Section 3 describes implementation of the algorithm as GFIT2. Section 4 presents tests of GFIT2, both for synthetic spectra and for real spectra taken to coincide with overpasses by aircraft making in situ CO_2 measurements. Section 5 presents preliminary conclusions to date, and Sect. 6 outlines future plans.

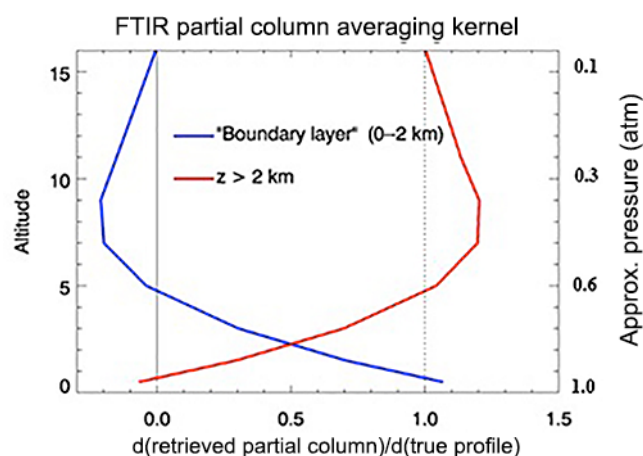


Figure 2. Partial column averaging kernels for the profile retrieval algorithm of Fig. 1 (GFIT2 applied to the 6220 cm^{-1} spectral band, with assumed signal-to-noise ratio = 1000).

2 Background and algorithm origins and history

GFIT is the algorithm adopted by the TCCON for analysis of the spectra; it was developed over many years by Geoff Toon at JPL. GFIT is also used to analyze MkIV balloon spectra (e.g., Sen et al., 1996) and was used in the Version 3 processing of ATMOS spectra (Irion et al., 2002). It is a profile scaling algorithm, employing a quasi-linear regression to derive scale factors for all important absorbers as well as other atmospheric and instrument parameters, such as continuum level and frequency shift.

GFIT is designed in such a way that its “forward model” is independent of and separable from its “inverse method”. These terms are discussed in Rodgers (2000), but briefly the forward model is an algorithm that calculates the atmospheric spectra comparable to the observed spectra, incorporating radiative transfer and molecular physics along with assumed gas distributions. The inverse method retrieves a state vector of parameters, such as molecular mixing ratio, by finding values which provide a best fit to the spectrum given other assumptions and constraints. The GFIT inverse method is a form of “optimal estimation” as described further below, which applies the Gauss–Newton method, iteratively estimating the parameters by successive approximation.

Ground-based spectra – at microwave, IR, and UV wavelengths – have been analyzed in selected applications for limited altitude profile information, for many years (Connor et al., 1995, 2007; Pougetchev et al., 1996; Schofield et al., 2004.) The physical origin of the limited profile information in ground-based spectra is somewhat varied. Most commonly, the pressure-broadened line shape is exploited; however the use of lines of varied opacity (see Fig. 1) and the use of multiple atmospheric paths (Schofield et al., 2004) are also sources of profile information.

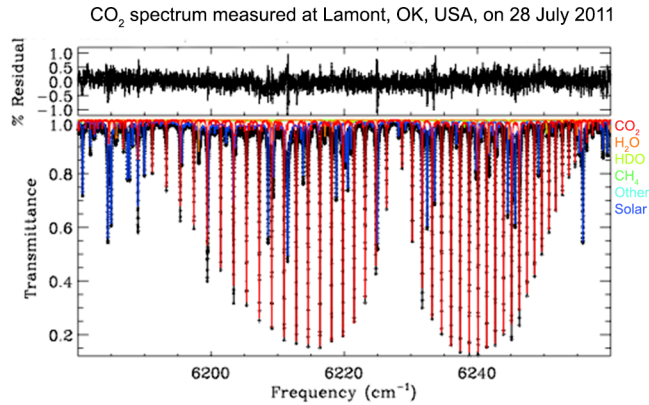


Figure 3. A typical spectrum in the 6220 cm^{-1} band. The calculated spectrum, produced by GFIT in profile scaling mode, is superimposed, with the individual gas contributions shown in color. CO_2 , solar lines, and H_2O dominate the visible features. The residuals (also typical) are shown in the upper panel. This spectrum is one of the atmospheric measurements used in Sect. 4.2 and subsequently.

The most common algorithm for the inverse method used in these and other studies is optimal estimation, formulated by Rodgers (1976, 2000). That algorithm will be described in detail in the following section.

Optimal estimation has been implemented as a user-selected option of inverse method added to the version of GFIT publically released in 2012; no other changes to the standard GFIT algorithm were made. The modified algorithm is known as GFIT2. GFIT is designed to treat each spectral band independently. All calculations in this paper are of the 1.61μ (6220 cm^{-1}) spectral band; the use of other bands will be discussed briefly in Sect. 5. Figure 3 shows a typical spectrum from the TCCON site at Lamont, OK, USA.

3 Algorithm and implementation

The optimal estimation formulation of Rodgers (2000) was adapted and applied for use with the “full physics” algorithm (inverse method plus forward model) developed for the first Orbiting Carbon Observatory (OCO) satellite, which failed to reach orbit in 2009. The OCO inverse method, as it existed in 2007, is described in Connor et al. (2008) and was used as a starting point for the development of GFIT2. Much of the discussion in Sect. 2 of Connor et al. (2008) is directly applicable.

3.1 Inverse method

The OCO inverse method was adapted for use with GFIT and is briefly described here. We use the notation and concepts of Rodgers (2000). The spectrum, or measurement vector \mathbf{y} , is expressed symbolically as $\mathbf{y} = F(\mathbf{x}) + \boldsymbol{\varepsilon}$, where \mathbf{x} is the

state vector, F is the forward model, and $\boldsymbol{\varepsilon}$ is the vector of measurement errors.

The solution of the GFIT2 inverse method is the state vector $\hat{\mathbf{x}}$ with maximum a posteriori probability, given the measurement \mathbf{y} . We solve for the state vector update $d\mathbf{x}_{i+1}$, using a slightly modified form of Rodgers’ Eq. (5.8), to improve numerical accuracy by avoiding the inversion of a large matrix:

$$\left(\mathbf{S}_a^{-1} + \mathbf{K}_i^T \mathbf{S}_\varepsilon^{-1} \mathbf{K}_i \right) d\mathbf{x}_{i+1} = \left[\mathbf{K}_i^T \mathbf{S}_\varepsilon^{-1} (\mathbf{y} - F(\mathbf{x}_i)) + \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a) \right], \quad (1)$$

where \mathbf{K} is the weighting function matrix, or Jacobian, $\mathbf{K} = \frac{\partial \mathbf{y}}{\partial \mathbf{x}}$; \mathbf{x}_a is the a priori state vector; \mathbf{S}_a is the a priori covariance matrix; and \mathbf{S}_ε is the measurement covariance matrix.

After each iteration, we test for convergence. To facilitate that, we compute the change in the solution scaled by its estimated variance:

$$d\sigma_i^2 = d\mathbf{x}_{i+1}^T \hat{\mathbf{S}}^{-1} d\mathbf{x}_{i+1}, \quad (2)$$

where $\hat{\mathbf{S}}$ denotes the covariance of the retrieved state, using the relation

$$\hat{\mathbf{S}}^{-1} d\mathbf{x}_{i+1} \cong \left[\mathbf{K}_i^T \mathbf{S}_\varepsilon^{-1} (\mathbf{y} - F(\mathbf{x}_i)) - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a) \right]. \quad (3)$$

$d\sigma_i^2$ is effectively the square of the state vector update in units of the solution variance.

If $d\sigma_i^2 < fn$ (where n is the number of state vector elements and f is an adjustable convergence parameter), convergence is reached.

Lastly, we compute the retrieval covariance matrix, $\hat{\mathbf{S}}$, and the averaging kernel matrix, \mathbf{A} . $\hat{\mathbf{S}}$ is given by

$$\hat{\mathbf{S}} = \left(\mathbf{K}^T \mathbf{S}_\varepsilon^{-1} \mathbf{K} + \mathbf{S}_a^{-1} \right)^{-1}. \quad (4)$$

The averaging kernel matrix \mathbf{A} is given by

$$\mathbf{A} = \hat{\mathbf{S}} \mathbf{K}^T \mathbf{S}_\varepsilon^{-1} \mathbf{K}. \quad (5)$$

Finally, the degrees of freedom (DoF) for signal are given by the trace of the matrix \mathbf{A} ; the degrees of freedom for the CO_2 profile are the trace of the CO_2 -only sub-matrix of \mathbf{A} .

To enable use of the Rodgers algorithm, a modified GFIT code was developed which completely separates the forward model and inverse method, and allows integration of optimal estimation profile retrieval with the existing code. Conceptually, the experimental, integrated GFIT allows selection of the existing (profile scaling) or modified (profile retrieval) algorithm. This is simply accomplished by setting a parameter in an input file. The integrated algorithm has input and output files identical to the existing GFIT, plus new input and output files specific to profile retrieval, which are not required unless the modified algorithm is selected.

3.2 Measurement error

A critical input to the algorithm of Eq. (1) is the measurement error covariance, S_e . It is assumed to be diagonal, and the simplest assumption is that the error is entirely random noise independent of frequency. As we will see later, in all real spectra there are systematic residuals, larger than the actual noise level, due to spectral features that can only be imperfectly modeled in the algorithm. If these features are not taken into account in constructing S_e , the retrieved profile develops severe oscillations.

The simplest way of “de-weighting” spectral features which remain in the residuals is to increase the estimated measurement error estimate (equivalently, reduce the assumed SNR) at all frequencies, so that residual features are ignored (treated as measurement error) (e.g., Connor et al., 1995). As we will see, in practice this is somewhat effective at damping oscillations, but only at the cost of losing most of the profile information in the spectrum.

An alternative approach we have attempted to avoid profile oscillations is to vary the assumed spectral error to reflect the real residuals obtained by spectral fits (Rodgers and Connor, 2003). A two-stage retrieval is run, in which stage 1 is profile scaling. The residuals from stage 1 are then used to estimate the spectral error as a function of frequency and inserted on the diagonal of S_e . This procedure greatly reduces profile oscillations in many synthetic retrievals. We refer to it as “variable SNR”.

A third approach to the problem of systematic residuals is to estimate them empirically, and then to include them in the forward model, multiplied by a scale factor retrieved as part of the state vector (JPL, 2015). Perhaps the simplest approach is to estimate the systematic component by averaging the residuals over the entire set of spectra under study. This technique and simple variants on it will be described in Sect. 4.4 below.

3.3 State vector and a priori uncertainties

The full state vector consists of the CO₂ profile, scale factors for the other gas profiles contributing to the spectrum in the band pass (H₂O, HDO, and CH₄ in the 6220 cm⁻¹ band); the background continuum level, tilt, and curvature; a frequency shift and a zero level offset. This is identical to the standard GFIT scale factor, except for the CO₂ profile itself. A scale factor multiplying a vector of systematic residuals, as described in Sect. 3.2, has been added to the state vector for the retrieval tests of Sect. 4.4.

A critical input is the a priori covariance matrix S_a , specifying assumed uncertainties in the state vector and their correlations. The retrievals in this paper assume that S_a is diagonal. The a priori uncertainties assumed are guided by those used in the standard GFIT scaling, namely 1 for the three interfering species and the continuum level, 0.1 for the continuum tile and curvature, 2 for frequency shift, and 0.5 %

for zero level offset (which is expected to be approximately zero). The uncertainty in each of the 70 levels in the CO₂ profile is set independently. These uncertainties range from 1 to 5 %, are largest near the surface, and have been adjusted to improve the test results where possible. Finally the residual scale factor, when in use, has been assigned an uncertainty of 10 %, based on the observed variability of the systematic residuals.

3.4 Other input parameters

The only other input parameters specific to profile retrieval concern convergence and goodness of fit. They include the convergence parameter defined in Sect. 3.1, the maximum acceptable X^2 of the spectral fit, and the maximum number of iterations allowed.

4 Testing GFIT2

4.1 Synthetic spectra

The algorithm was first tested by retrievals on synthetic spectra, where the “true” atmospheric profile is known. In these tests, the forward model (used by the algorithm) may be the same as the forward function (which includes all true physics) or may differ from it in a controlled way.

4.1.1 No forward-model error

We illustrate the most basic test in Fig. 4. Here a synthetic spectrum was calculated from the profile labeled true (diamonds); then GFIT2 was run on the calculated spectrum without modification, using the a priori profile shown. (The a priori profile is selected on the basis of climatology.)

The assumed signal-to-noise ratio was 1000. The solid lines in Fig. 4 show the retrieved profile, and the degrees of freedom for the profile are shown in the legend. The two results shown in Fig. 4a and b are typical. The retrieved profile has 3.3–3.5° of freedom and follows the departure of the true profile from the a priori reasonably well. This behavior is consistent with previous experience and with expectations, and it leads to the conclusion that the algorithm is working as designed.

4.1.2 Pointing error

Next we used a known instrumental limitation to test the skill of the variable-SNR technique. Namely, while the instrument nominally points at the center of the solar disk, it is common for some error to be introduced by patchy cloud cover or simply by tracking hardware problems. The effect of such pointing error is to introduce a Doppler shift due to solar rotation, making calculations of the solar Doppler shift inaccurate, and thus result in an uncompensated shift in the position of solar lines relative to telluric lines.

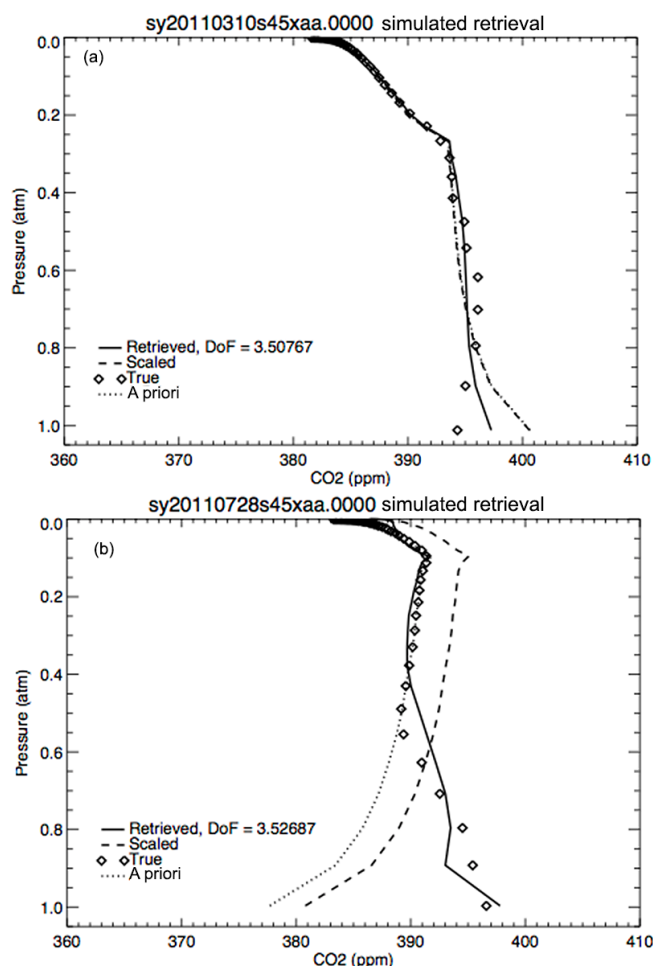


Figure 4. (a) Retrievals from a synthetic spectrum with no forward-model error. (b) As in (a) but assuming different true and a priori profiles.

To assess the effect of pointing error, we assumed that an error was present equal to 10 % of the solar diameter, which produces an error in the solar Doppler shift of ~ 1.3 ppm. We then applied the variable-SNR modification to the measurement error covariance \mathbf{S}_ϵ , as described in Sect. 3.2. Retrievals with this assumed error are shown in Fig. 5. There is a significant decrease in the degrees of freedom, from ~ 3.3 – 3.5 to ~ 2.6 . This occurs because the apparent SNR decreases in the regions of the solar lines. That is, the assumed measurement error is increased at all frequencies where the Doppler shift error causes an increased residual spectrum, and in some cases the solar lines and CO_2 features overlap; the increased measurement error reduces the algorithm's sensitivity to the measured spectrum, corresponding to lower DoF.

The retrieval in Fig. 5a is qualitatively similar to the one with higher DoF in Fig. 4a. The retrieved profile in Fig. 5b spreads the increased CO_2 in the lower troposphere over a broad range of altitude, showing the effect of poorer vertical sensitivity compared to Fig. 4b, but does detect the presence

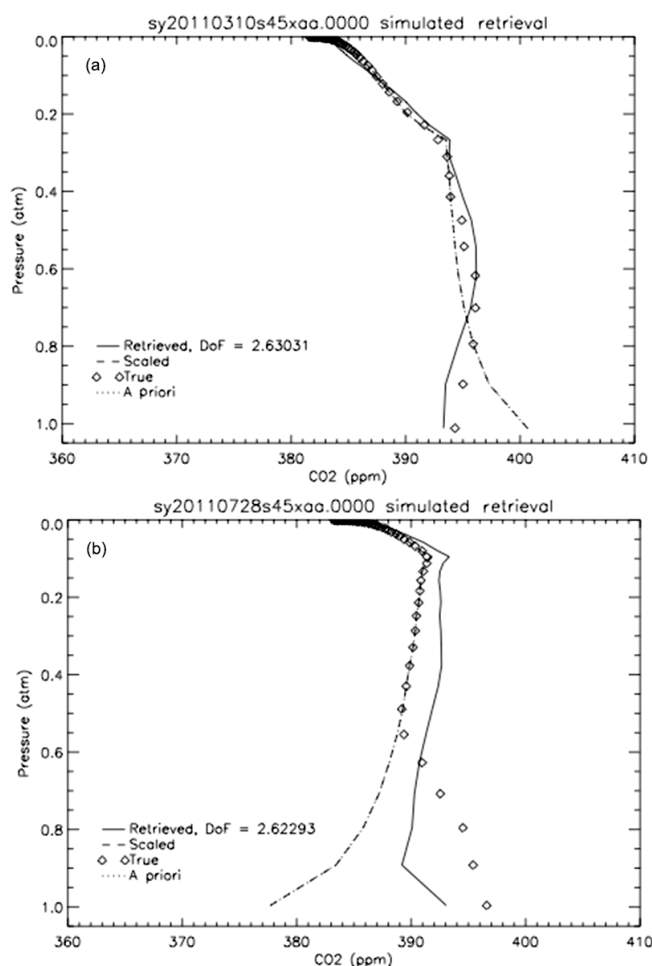


Figure 5. (a) As Fig. 4a, with telescope pointing error assumed. (b) As Fig. 4b but with telescope pointing error assumed.

of enhanced CO_2 in the troposphere. Overall, we believe the variable-SNR modification is shown to be effective in coping with systematic residuals due to the level of solar pointing error introduced.

4.1.3 Linewidth error

Error in spectroscopic parameters is an important cause of systematic error in the calculated spectra, leading to systematic structures in the spectral residuals. Since the profile retrieval depends critically on the spectral line shape, spectroscopic errors will limit its performance, possibly severely. The most obvious source of spectroscopic error affecting line shape is the pressure-broadening coefficient, which simply scales the linewidth at a given pressure. It is arguably the largest source of line shape error as well. We will use synthetic spectra and simulated retrievals to evaluate its effects.

For this purpose we have multiplied the pressure-broadening coefficients in the relevant CO_2 bands near $1.6 \mu\text{m}$ by 1.01, thus modeling a 1 % error in linewidth, and

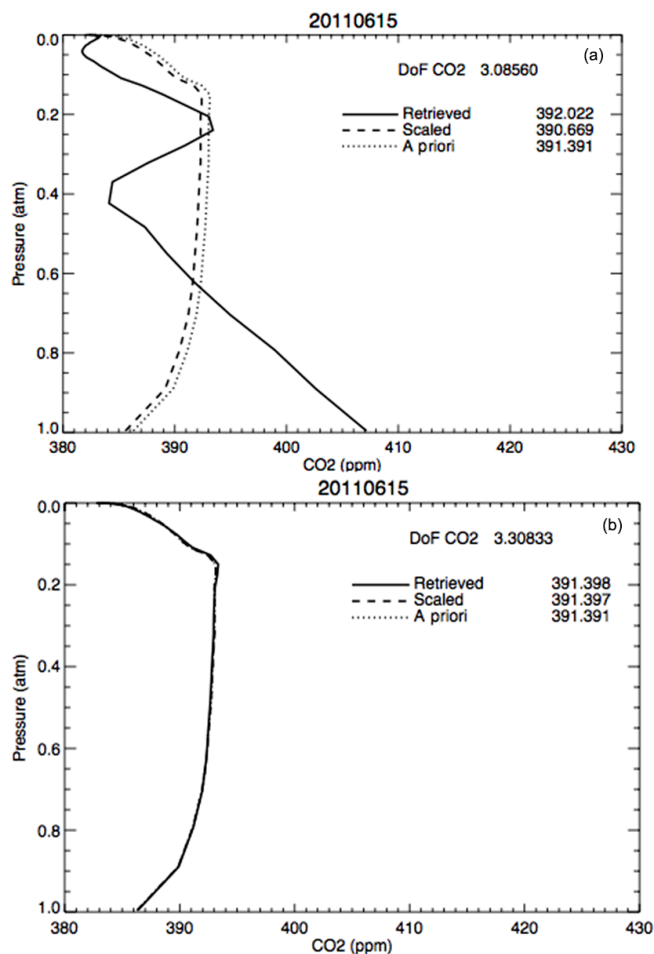


Figure 6. (a) Retrievals with a 1 % error in pressure-broadening coefficient assumed. (b) As (a) except the residuals of the scaling retrieval are included in the profile retrieval's forward model.

used these modified coefficients to calculate sets of ~ 100 synthetic spectra on specific days at Lamont, using actual SZA and modeled temperature, etc. We have then run both the scaling and profile retrievals with the original unmodified coefficients. The average profiles for 1 day are shown in Fig. 6a.

The scaling retrieval reduces the CO_2 mixing ratio slightly to compensate for the 1 % linewidth error, producing a net error in X_{CO_2} (and mixing ratio) of 0.2 %. The profile retrieval, on the other hand, produces large oscillations, of ~ 5 % at the surface and ~ 2 % in the upper troposphere and stratosphere. Despite these large errors in profile, the net error in X_{CO_2} is of similar magnitude (but opposite sign) to the scaling retrieval.

Fortunately, errors in real retrievals are unlikely to be as large as in these simulations.

Much effort in recent years has gone into refining knowledge of the spectroscopic parameters needed for modeling atmospheric CO_2 . Devi et al. (2007) state typical uncertainty

in the pressure-broadening coefficient of strong lines in the $1.6 \mu\text{m}$ CO_2 bands is only approximately 0.1 %. Admittedly, this is a formal uncertainty derived from their spectral fits and may not include some sources of absolute uncertainty. However even an absolute uncertainty that small would produce an error in CO_2 at the surface of ~ 2 ppm.

The extreme sensitivity of the CO_2 profile to errors in the pressure-broadening coefficient, and by extension to other sources of line shape error, motivates a search for ways to “correct” the forward-model spectra to minimize such effects. Since the (unknown) true error in a spectroscopic parameter is constant, it may be expected to produce a spectral signature which is very similar from measurement to measurement, over many measurements. If we can isolate and remove that signature, the profile retrieval may be able to capture variations in profile shape within that set of measurements.

As a preliminary test, we assume that the spectroscopic error signature is given by the residuals of the scaling retrieval and add those to the calculated spectrum in the profile retrieval. The average profiles, corresponding to Fig. 6a, are shown in Fig. 6b.

The average retrieved profile is nearly identical to the scaling retrieval; no spurious changes in profile shape are introduced. The derived X_{CO_2} mole fractions differ by only 0.01 ppm.

Of course, for real measurements, the signature of spectral error is not so easily derived. Later, in Sect. 4.4, we calculate the mean residual vector for large sets of real measurements and attempt profile retrievals including a scale factor applied to the mean residual vector.

4.1.4 ILS error

Another potentially significant source of error is distortion of the measured line shape itself. For Fourier transform spectrometers (FTSs, as used by TCCON) the ILS is a convolution of contributions from the finite path difference and the finite field of view (FOV) of the FTSs. The path difference and its ILS contribution (a sinc function) are well known, but the FOV, which contributes a rectangular shape, has an uncertainty we estimate as 7 %. This causes the observed line to be broader and weaker than the atmospheric line, and it progressively has a larger effect as the line is narrower; i.e., the error due to finite aperture becomes more important at lower pressure where the intrinsic line shape is narrower (see for example Davis et al., 2001).

We illustrate this effect in Fig. 7, which is calculated for the same spectra as Fig. 6 and so is directly comparable to Fig. 6a. The net effect of this error is very small in the lower troposphere and grows only to ~ 1 % in the stratosphere. We conclude that error in the measured line shape is unlikely to dominate error in the calculated line shape (Sect. 4.1.3).

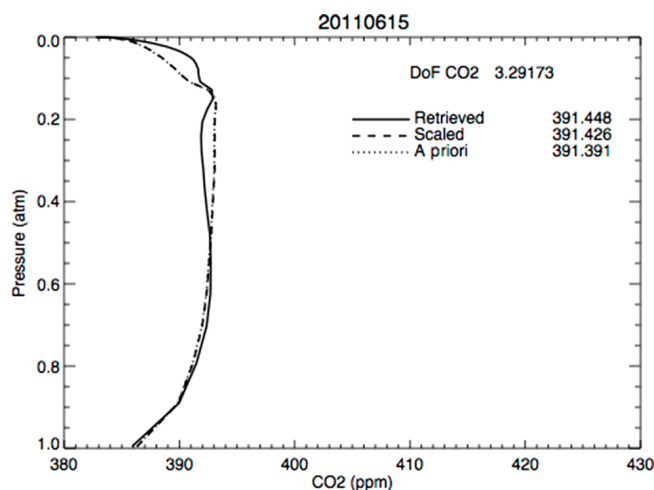


Figure 7. Retrievals of the spectra used for Fig. 6 but with an assumed error of 7 % in the instrument field-of-view.

4.2 Atmospheric measurements

Atmospheric spectra are routinely measured in Lamont, Oklahoma, at the Southern Great Plains site of the Department of Energy Atmospheric Radiation Measurement network. A Cessna aircraft equipped with air sampling in situ detectors is flown there on a regular basis and produces CO₂ profiles from ~ 0 to 5 km altitude. (Biraud et al., 2013). Although these profiles include only about half the total column of atmospheric CO₂, it is in the lowest few kilometers that CO₂ is most variable and least predictable. Therefore the Cessna measurements, coupled with climatological estimates at higher altitudes, are expected to produce reasonable estimates of the full CO₂ profile. In particular, they can test the profile retrieval algorithm's ability to detect variations from climatology in the lower troposphere.

With this in mind, we have chosen several days for study. On these days, Cessna flights were made, atmospheric conditions were excellent, and many high-quality near-infrared spectra were recorded. We selected days from various times of year to allow conditions as variable as possible. The specific days chosen are 15 June 2011, 5 July 2011, 28 July 2011, 26 August 2011, 24 December 2011, 14 January 2012, and 15 January 2012.

Analysis of the data from these days immediately revealed significant errors in the solar Doppler shift. This is not only shown by simple examination of the residuals but is also formally calculated, as the difference in the frequency shift observed for solar lines (after correcting for the calculated Doppler shift) and telluric lines. GFIT does not automatically take this error into account, by recalculating the spectrum with the correct Doppler shift. However, these errors can be corrected by using the retrieved solar–telluric difference to correct the calculated Doppler shift, and then re-running the retrieval. All measured spectra and retrievals used and/or

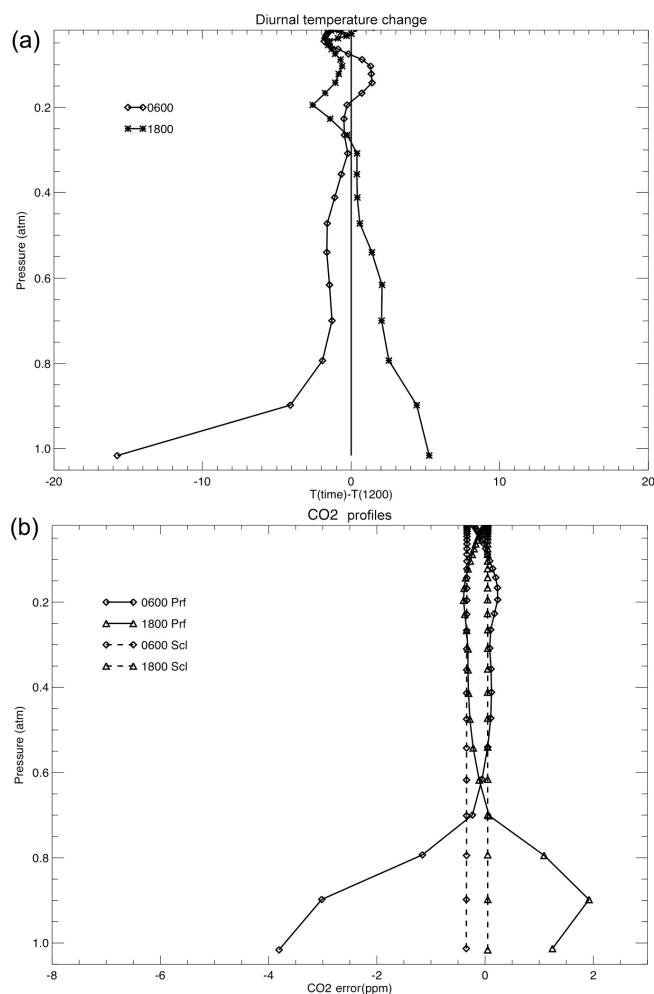


Figure 8. (a) NCEP temperature profiles at 06:00 and 18:00 for a single day, less the profile at 12:00 on the same day. (b) Error in CO₂ retrievals produced by not accounting for the temperature changes of (a). Dashed lines for profile scaling, solid for profile retrieval.

shown in this paper have been “Doppler-shift-corrected” in this way.

Another potential issue is the temperature profile used in performing each retrieval.

In clear, dry conditions, the temperature at the surface will sometimes vary by as much as 20 K during daylight. This has an important impact on retrieval of tropospheric CO₂, as shown in Fig. 8. Figure 8a shows the difference in temperature between 12:00 LMST and 06:00 and 18:00 of the same day. Figure 8b shows the effect of not correctly including this temperature variation. In particular Fig. 8b shows the difference between profiles retrieved assuming the 12:00 temperature profile and those using the temperatures of 06:00 and 18:00. Two sets of curves are shown, one for profile scaling and one for profile retrieval.

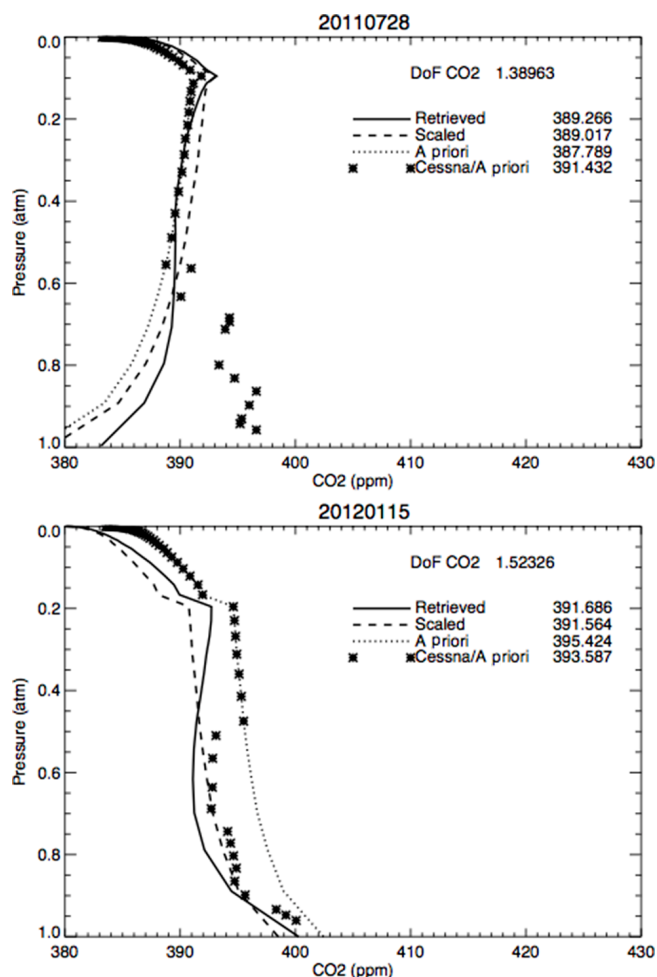


Figure 9. (a) Average retrieval within about 1 h of a Cessna overflight, assuming SNR = 100, on 28 July 2011. (b) As (a) but for 15 January 2012.

Note that the CO₂ errors produced by imposing an error in surface temperature are largely confined to the lowest 2 km. To minimize this effect, we have used the NCEP profiles at 06:00, 12:00, and 18:00 LMT, and interpolated between them to the approximate time of the Cessna overflights. These interpolated temperature profiles have been used in all the retrievals shown subsequently in this section.

4.2.1 SNR = 100

As described in Sect. 3.2, we first attempted retrievals by setting the SNR low enough to avoid trying to fit systematic spectral residuals. The SNR observed to achieve this for the current dataset is approximately 100. Two examples are shown in Fig. 9. Figure 9a may be compared directly to Figs. 4b and 5b. A smoother version of the Cessna data is the true profile in Figs. 4b and 5b. The a priori profile is the same. Note that the degrees of freedom for the CO₂ profile (“DoF CO₂”) are 1.4–1.5, implying there is only at best slightly

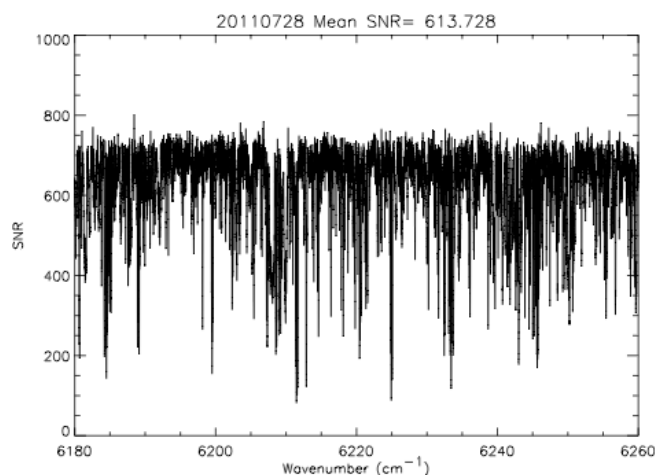


Figure 10. Signal-to-noise ratio (SNR) derived for 28 July 2011 from the variable-SNR technique (see text).

more information about the profile than the 1° of freedom available to the scaling retrieval.

We see that the retrieval in Fig. 9a poorly matches the Cessna profile in the lower troposphere, although it is a small improvement on the a priori. The profile in Fig. 9b matches CO₂ at $p > 0.5$ atm reasonably well, though that may be fortuitous.

For present purposes the thing to note is that assuming SNR = 100 not only largely masks information on the altitude profile but also avoids profile oscillations (e.g., Fig. 6a), at least for these 2 days.

4.2.2 Variable SNR

We attempted to include as much profile information as possible while avoiding “over-fitting” the spectral regions with the poorest residuals, by using the variable SNR as described in Sect. 3.2. We illustrate by showing results for the same days as in the preceding section, 28 July 2011 and 15 January 2012. We show in Fig. 10 the effective SNR on which we based the diagonal of S_e for use in Eq. (1), for the day 28 July 2011.

Note that the effective SNR varies from ~ 100 to 750. The mean value is shown in the title and is a bit over 610. This is to be compared to SNR = 1000 in Fig. 4 and SNR = 100 in Fig. 9.

Examples of profiles retrieved with variable SNR are given in Fig. 11. CO₂ DoF has increased to ~ 2.9 –3.3, compared to Fig. 9. Figure 11a is directly comparable to Fig. 9a. There appears to be some improvement in the lower troposphere but degradation in the upper troposphere, with a suggestion of an incipient oscillation. Figure 11b shows a runaway oscillation. This result shows that the fundamental instability of the retrieval has not been adequately mitigated by the application of the variable-SNR technique. It is worth pointing out that these 2 days are “typical” of the 7 days studied. In par-

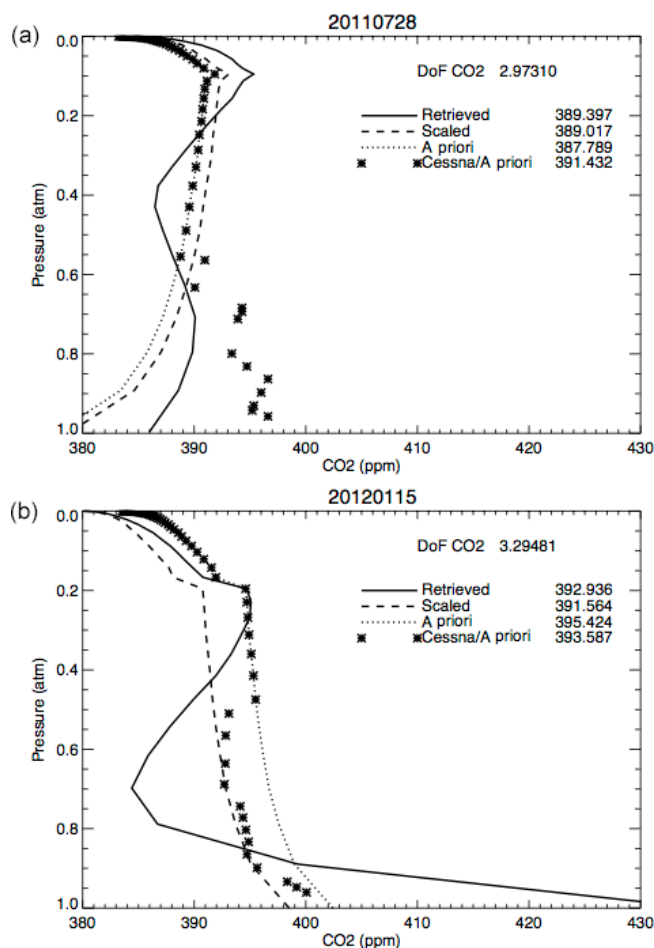


Figure 11. (a) Average retrieval within about 1 h of a Cessna over-flight, assuming variable SNR, on 28 July 2011. (b) As (a) but for 15 January 2012.

ticular, the 2 other winter days show oscillations much like in Fig. 11b; the 3 other summer days do not, but they show little if any improvement vs. the Cessna profile.

The performance of the profile retrieval algorithm with variable SNR on the measured spectra tested is clearly unsatisfactory. In an effort to understand this limitation we next closely examine the spectral residuals the algorithm produces, as described in the next section.

4.3 Spectroscopic residuals

Past experience indicates that the oscillatory behavior of the profiles seen in Fig. 11 is most likely driven by a failure of the forward model to adequately reproduce the measured spectrum. This of course may reflect systematic error in either the model or the instrument. To isolate the spectral signature of the error, we calculated mean residuals from many spectra. We further expanded our study to examine the residuals produced by the same model instrument at a different site, namely Lauder, New Zealand.

On the 7 days examined in Sect. 4.2, a total of 5946 good-quality spectra were recorded at Lamont. There were ~ 600 spectra day^{-1} in winter and up to ~ 850 day^{-1} in summer. We have performed retrievals on all of these spectra and calculated the mean residual for each day, along with the overall mean residual, in an attempt to isolate systematic spectral features which the GFIT forward model cannot reproduce.

At Lauder we examined every 10th spectrum recorded on 13 days selected for clear sky and seasonal coverage. The days were 10 July 2010, 11 July 2010, 30 July 2010, 24 August 2010, 8 September 2010, 2 November 2010, 7 November 2010, 8 November 2010, 4 February 2011, 16 February 2011, 1 April 2011, 21 May 2011, and 28 September 2011. A total of 621 spectra were included; the number of spectra per day ranged from ~ 10 to 70.

Figures 12–13 show expanded views of portions of the mean residuals for the two sites, for air mass < 2 (Fig. 12) and for $2 < \text{air mass} < 4$ (Fig. 13). A pair of two (upper and lower) panels is shown for each spectral interval, 6205–6210 and 6240–6245 cm^{-1} . The upper panel shows the mean residual. The dashed vertical lines indicate the positions of CO_2 spectral lines. The lower panel shows the standard deviation of the daily mean residual vectors.

It is immediately clear that there are systematic residuals of $\sim 0.5\%$ depth at all of the CO_2 positions. Also, the CO_2 residuals are very similar at Lauder and Lamont, and for both ranges of air mass. Closer examination shows that the residuals are slightly asymmetric, such that the line center is at slightly higher frequency than the center of the residual.

While to first order all the CO_2 residuals in Figs. 12–13 are very similar, there are also differences. The systematic residuals are broader at higher air mass, and the same features are somewhat broader at Lauder than at Lamont.

4.4 Mean bias correction

We strongly suspect that a stable profile retrieval is not possible in the presence of systematic spectral errors as suggested by the residuals of Sect. 4.3 and that these will readily produce the unsatisfactory oscillations seen in Fig. 11. This systematic spectral signature might be thought of as a “bias” of the GFIT forward model which prevents it from fitting the measured spectra precisely enough. The GFIT forward model, as with most practical atmospheric spectral line models, uses the Voigt profile as its line shape. However, it has recently become well understood that the Voigt line shape is inadequate to model atmospheric spectra at the sub-percent level. See, for example, Fig. 1 of Long et al. (2011). Unfortunately, improved line shape functions are far more complex, and, while several of them are known to improve spectral fits in the laboratory (ibid.), there is no agreement as to which of them contains the best physical description of the line formation. Since atmospheric spectra are formed in far different physical conditions than laboratory spectra, it is unclear how

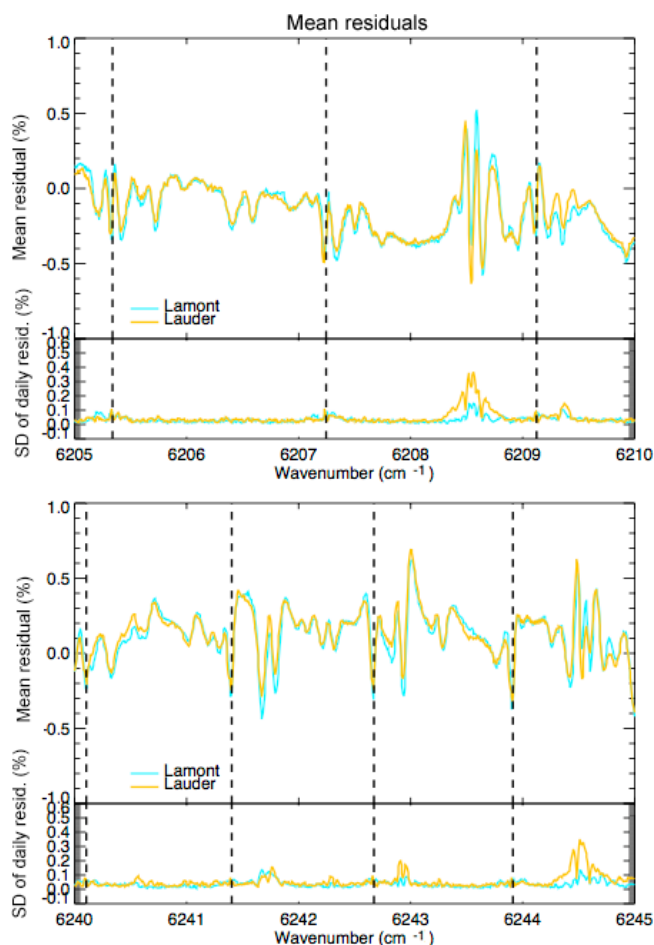


Figure 12. Mean and standard deviation of residuals in selected spectral intervals, for measurements with air mass < 2 .

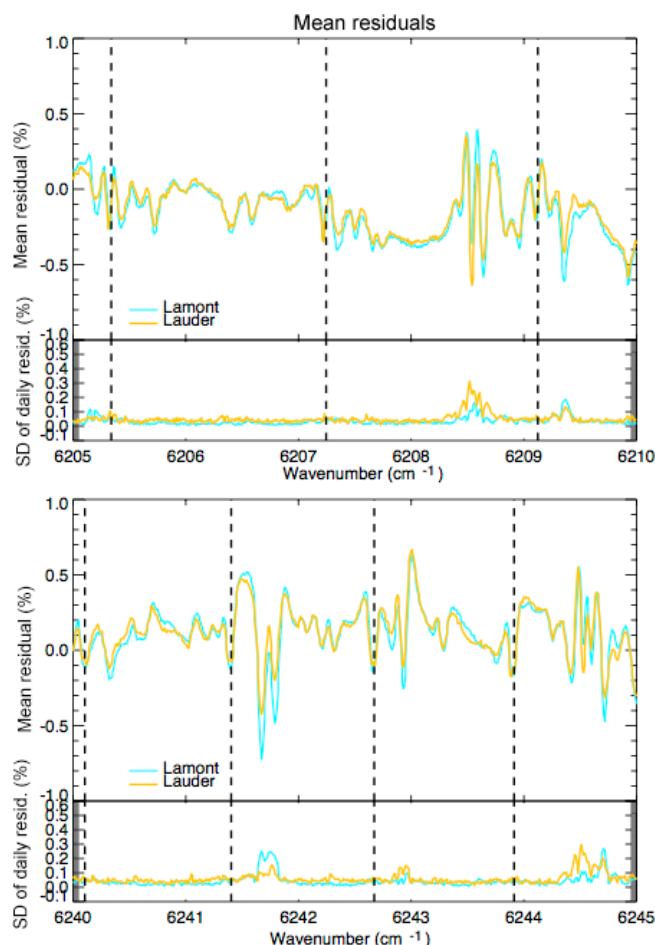


Figure 13. Mean and standard deviation of residuals in selected spectral intervals, for measurements with $2 < \text{air mass} < 4$.

to modify the forward model to improve the observed spectral fits.

Pending a future clarification of the physics of line formation, we have attempted to stabilize the algorithm by performing simple “corrections” to the forward model to remove, as far as possible, the spectral bias. In the first instance, we have performed retrievals on the Lamont spectra discussed in Sect. 4.2, accounting for systematic spectral residuals as follows. We have modified the forward model to include addition of a spectral basis vector, multiplied by a scale factor, to the modeled spectrum. We calculate the mean residual spectrum from a large set of the Lamont retrievals (the set to be defined shortly). We then use those mean residuals as the basis vector to be added to the modeled spectrum. The scale factor which multiplies the basis vector is incorporated in the state vector, to be retrieved for each measured spectrum. It is typically ~ 1 .

In the first instance, we derive the mean residual from the full set of 7 days of data and use this as the basis vector in retrievals from each measured spectrum. We show two of the

daily retrieved profiles (each the average of ~ 80 individual retrievals) in Fig. 14.

Figure 14a and b are directly comparable to Fig. 11a and b. The same data and algorithm are used except for the addition of the scaled mean residual as just discussed.

The comparison of Figs. 11 and 14 shows a dramatic improvement on 15 January 2012, eliminating the large oscillation of Fig. 11b. Also on 28 July 2011, the profile of Fig. 14a, after subtracting the mean residual, is less oscillatory than previously (Fig. 11a). On both days, the integrated column from the profile retrieval, represented by XCO_2 , is slightly closer to the estimated true value. This is promising and suggests further development of the “bias correction” procedure.

Deriving the residual vector from the full 7 days of measurements implicitly assumes that the residuals are independent of seasonal effects and instrumental adjustments over a long period. This is unlikely to be the case; in fact we have already noted in discussing Figs. 12–13 that the residuals depend to some degree on air mass. The mean air mass of a set of spectra will vary with season and times of day. To lessen

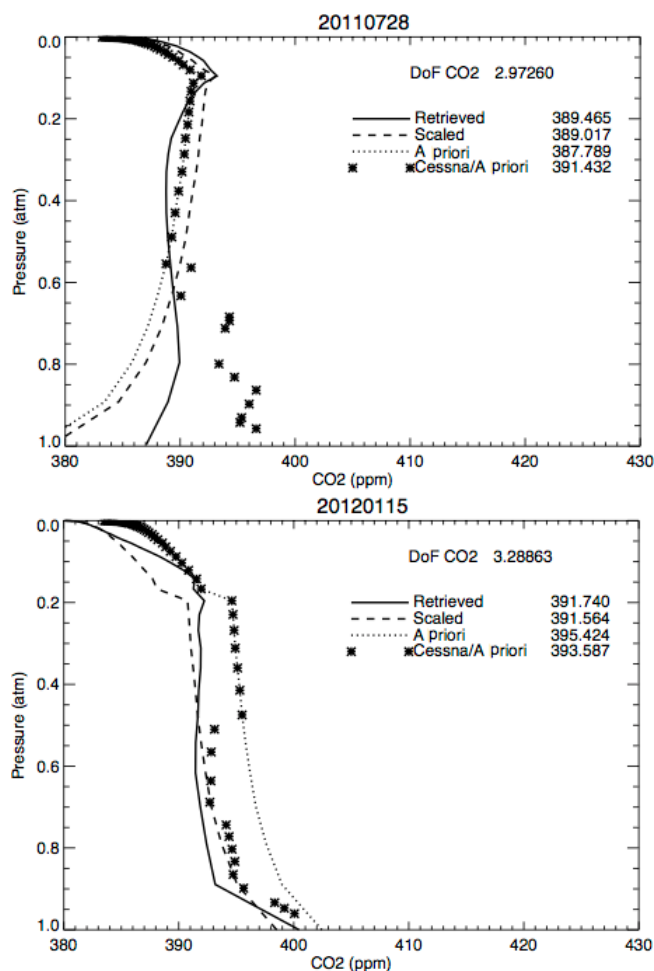


Figure 14. (a) As Fig. 11a but including a basis vector of mean residuals in the forward model. (b) As (a) but for 15 January 2012.

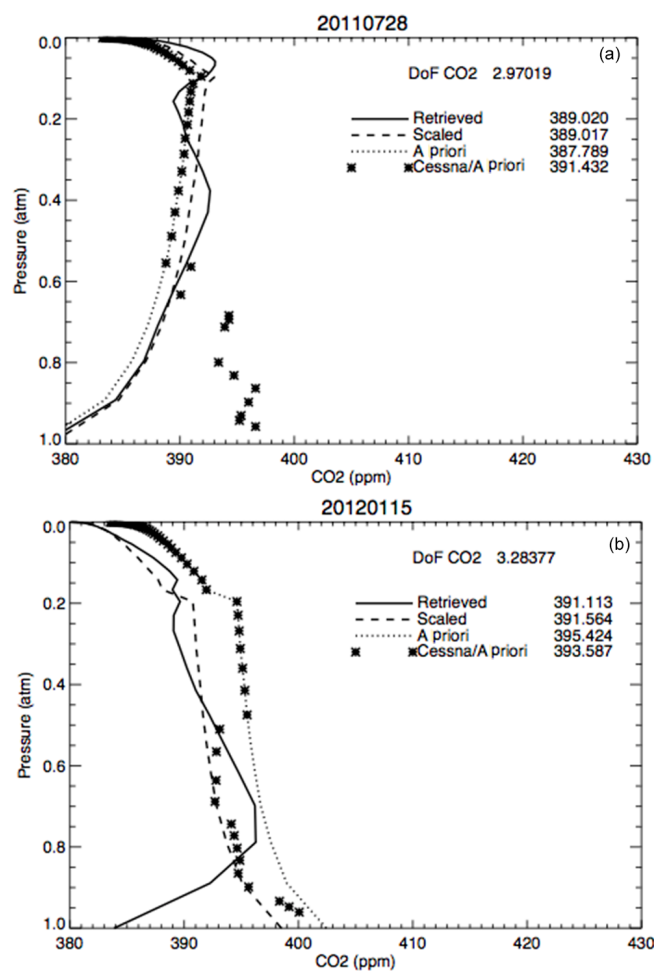


Figure 15. (a) As Fig. 14a but with a different residual basis vector. (b) As Fig. 14b but with a different residual basis vector.

the impact of both air mass and potential instrumental variations in the residuals, we have used monthly residuals calculated in a limited range of air mass (1–2 or 2–4) as the spectral correction vectors and run the retrievals once more. The results are shown in Fig. 15.

Unfortunately these results are a clear step backward. Figure 15a shows no sensitivity to the enhanced lower-tropospheric CO₂, and in Fig. 15b we see the return of oscillatory behavior. The XCO₂ value from the profile retrieval minus the estimated true value is similar to the scaled minus true value on 28 July 2011, and somewhat larger than scaled minus true value on 15 January 2012. It should be noted that the final residuals resulting from addition of the scaled, monthly mean residuals are better (the rms fit is smaller) than those which result from use of the overall mean residuals. Despite that, the profiles are worse, suggesting that spectral features relevant to the CO₂ profile are being removed by the attempt at spectral bias correction.

Our best results to date come from adding the scaled mean residuals of the set of days under study. With that in mind we

will expand the discussion to include days other than the two illustrations used so far.

Figure 16a and b show results for 26 August 2010 and 14 January 2012. They were produced in the same way as Fig. 14a and b, that is, including addition of a scaled mean residual. The two dates are chosen to illustrate two features of the full set of retrievals. Namely, on 26 August we see a smooth profile with only a suggestion of oscillation, which seems (maybe fortuitously) to track some enhanced CO₂ in the lower troposphere. This description is similar to one for 28 July (Fig. 14a); in fact it is typical of all 4 summer days of this group. Conversely, 14 January shows a serious oscillation in the profile, unlike 15 January (Fig. 14b). The third winter day studied, 24 December 2011 (not shown), has an oscillation similar to 14 January. On both days, XCO₂ from the profile retrieval is slightly closer to the true value than is the scaled value.

In summary these results fall into two classes. In one, the retrieved profile is reasonably well behaved but offers little if any improvement on the profile scaling version. In the

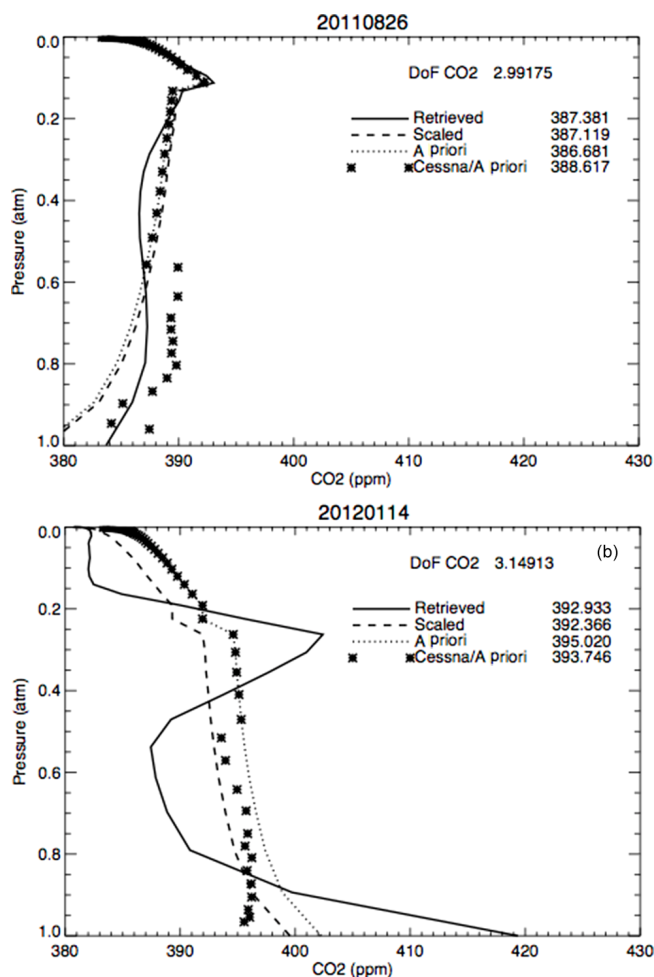


Figure 16. (a) As Fig. 14a but for 28 August 2011. (b) As Fig. 14b but for 14 January 2012.

other, the retrieved profile suffers serious oscillations. X_{CO_2} from the profile retrieval is similar to that from the scaling retrieval.

5 Conclusions

The algorithm behaves as expected on synthetic data. On real data, results are usually worse than scaling, given our a priori knowledge of the CO_2 distribution with altitude. Spectral residuals are generally poor. When modifications to the spectra and/or tight constraints force residuals to be small, profile oscillations tend to be severe. Based on the tests shown in Sect. 4.1, it would seem that our theoretical knowledge of the atmospheric spectra is inadequate to provide useful CO_2 profiles at the accuracy required. It is important to consider the word “useful”: the required accuracy for useful measurements of CO_2 in the troposphere is very high (~ 0.1 – 0.2%) relative to other atmospheric species and altitude regions.

Demands on our knowledge of the spectra are correspondingly high.

There are at least two directions to follow in pursuit of useful profile retrievals. One is improvements to the forward model. These could be in the form of more accurate values of spectral parameters, more appropriate models of spectral line shape, and/or knowledge of the instrument line shape. All of these areas have, however, already been the focus of intense work over an extended period of time, and breakthroughs may be slow in coming.

A second alternative is to exploit profile information from sources other than the pressure-broadened line shape. An immediately accessible source is spectral regions of higher and lower opacity than the spectral band considered here. In particular, several other CO_2 bands of varying opacity are routinely measured simultaneously with the $1.61\text{ }\mu\text{m}$ (6220 cm^{-1}) band by the TCCON FTSs. A profile retrieval using several bands simultaneously should be explored. For example, the $2.06\text{ }\mu\text{m}$ (4852 cm^{-1}) band has much higher opacity, while the $1.65\text{ }\mu\text{m}$ (6073 cm^{-1}) band has lower opacity than the $1.61\text{ }\mu\text{m}$ band we have used. As shown in Fig. 1, regions of high opacity are very sensitive to the lower troposphere, while more optically thin regions are sensitive to the upper troposphere and stratosphere. Simultaneous retrievals using all three bands would be far less sensitive to details of the spectral line shape and thus might avoid the difficulties described earlier. A complicating factor, however, is the likelihood of different errors in band strength in the three regions. A strategy for self-consistent scaling the band strengths might be required before performing a simultaneous profile retrieval.

An apparently simple alternative has been suggested, namely imposing a priori constraints on the profile shape by experimenting with explicit interlayer correlations in the a priori covariance matrix \mathbf{S}_a . Such correlations are not necessary simply to preserve profile smoothness. This is fundamentally because in the Rodgers algorithm the a priori profile shape implicitly imposes the profile fine structure, which is not strongly influenced by the measurement (see Rodgers, 1990). So, for example, the algorithm developed at Stony Brook University for ground-based microwave measurements of ClO has been used successfully for more than 20 years and has always used a diagonal \mathbf{S}_a matrix (Solomon et al., 2000; Connor et al., 2013).

Nevertheless explicit interlayer correlations may damp undesirable oscillations, and their effect should be explored. They have been used routinely in the OCO/OCO-2 retrieval algorithms (Connor et al., 2008; JPL, 2015), where experience shows that the nature and strength of correlations is key to doing this successfully. A whole series of experiments, analogous to those presented in Sect. 4 of this paper, could be envisioned to decide how best to apply correlations and to evaluate their efficacy. This would be a valuable part of a follow-on study, especially if combined with the multiband

retrieval approach. Regrettably, it is beyond the scope of the present effort.

Finally, it is our intention to release GFIT2 to the community, as an option within the public version of GFIT. That would allow testing and development by a wider range of experienced researchers. So far that has proven impractical, but we hope to do so in the near future.

Acknowledgements. Part of this research was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. We thank NASA's Carbon Cycle Science Investigation Program for supporting the development of GFIT2 (NNX14AI60G). Operations of TCCON at Lamont, Oklahoma, are made possible by NASA's OCO-2 project in collaboration with the DOE ARM program. Cessna data from the SGP are available through the ARM archive (www.archive.arm.gov). We thank Sebastien Biraud for his assistance in interpreting the aircraft data.

Edited by: F. Hase

Reviewed by: two anonymous referees

References

- Biraud, S. C., Torn, M. S., Smith, J. R., Sweeney, C., Riley, W. J., and Tans, P. P.: A multi-year record of airborne CO₂ observations in the US Southern Great Plains, *Atmos. Meas. Tech.*, 6, 751–763, doi:10.5194/amt-6-751-2013, 2013.
- Connor, B. J., Siskind, D. E., Tsou, J. J., Parrish, A., and Remsberg, A. E. E.: Ground-based microwave observations of ozone in the upper stratosphere and mesosphere, *J. Geophysical Res.*, 99, 16757–16770, 1994.
- Connor, B. J., Parrish, A., Tsou, J.-J., and McCormick, M. P.: Error analysis for the ground-based microwave ozone measurements during STOIC, *J. Geophys. Res.*, 100, 9283–9292, 1995.
- Connor, B. J., Mooney, T., Barrett, J., Solomon, P., Parrish, A., and Santee, M.: Comparison of ClO measurements from the Aura Microwave Limb Sounder to ground-based microwave measurements at Scott Base, Antarctica, in spring 2005, *J. Geophys. Res.*, 112, D24S42, doi:10.1029/2007JD008792, 2007.
- Connor, B. J., Bösch, H., Toon, G., Sen, B., Miller, C., and Crisp, D.: Orbiting Carbon Observatory: Inverse method and prospective error analysis, *J. Geophys. Res.*, 113, D05305, doi:10.1029/2006JD008336, 2008.
- Davis, S. P., Abrams, M. C., and Brault, J. W.: *Fourier Transform Spectroscopy*, Academic Press, 2001.
- Devi, V. M., Benner, D. C., Brown, L. R., Miller, C. E., and Toth, R. A.: Line Mixing and Speed Dependence in CO₂ at 6227.9 cm⁻¹: Constrained multispectrum analysis of intensities and line shapes in the 30013–00001 band, *J. Mol. Spec.*, 245, 52–80, 2007.
- Dohe, S.: Measurements of atmospheric CO₂ columns using ground-based FTIR spectra Doctor of Science dissertation, Karlsruhe Institute of Technology, 101–104, 2013.
- Irion, F. W., Gunson, M. R., Toon, G. C., Chang, A. Y., Eldering, A., Mahieu, E., Manney, G. L., Michelsen, H. A., Moyer, E. J., Newchurch, M. J., Osterman, G. B., Rinsland, C. P., Salawitch, R. J., Sen, B., Yung, Y. L., and Zander, R.: Atmospheric Trace Molecule Spectroscopy (ATMOS) Experiment Version 3 data retrievals, *Appl. Opt.*, 41, 6968–6979, 2002.
- JPL: Jet Propulsion Laboratory, Orbiting Carbon Observatory – 2, Level 2 Full Physics Algorithm Theoretical Basis Document, http://disc.sci.gsfc.nasa.gov/OCO-2/documentation/oco-2-v6/OCO2_L2_ATBD.V6.pdf, 2015.
- Kuai, L., Wunch, D., Shia, R.-L., Connor, B., Miller, C., and Yung, Y.: Vertically constrained CO₂ retrievals from TCCON measurements, *J. Quant. Spec. Rad. Trans.*, 113, 1753–1761, 2012.
- Long, D. A., Bielska, K., Lisak, D., Havey, D. K., Okumura, M., Miller, C. E., and Hodges, J. T.: The air-broadened, near-infrared CO₂ line shape in the spectrally isolated regime: Evidence of simultaneous Dicke narrowing and speed dependence, *J. Chem. Phys.*, 135, 064308, doi:10.1063/1.3624527, 2011.
- Pougatchev, N. S., Connor, B. J., Jones, N. B., and Rinsland, C. P.: Validation of ozone profile retrieval from infrared ground-based solar spectra, *Geophys. Res. Lett.*, 23, 1637–1640, 1996.
- Rodgers, C. D.: Retrieval of Atmospheric Temperature and Composition From Remote Measurements of Thermal Radiation, *Rev. Geophys. Space Phys.*, 14, 609–624, 1976.
- Rodgers, C. D.: *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific Publishing Co. Ltd., 2000.
- Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, *J. Geophys. Res.*, 108, 4116, doi:10.1029/2002JD002299, 2003.
- Schofield, R., Connor, B. J., Kreher, K., Johnston, P. V., and Rodgers, C. D.: The retrieval of profile and chemical information from ground-based UV-Visible DOAS measurements, *J. Quant. Spectros. Ra.*, 86, 115–131, 2004.
- Sen, B., Toon, G. C., Blavier, J.-F., Fleming, E. L., and Jackman, C. H.: Balloon-borne observations of mid-latitude fluorine abundance, *J. Geophys. Res.*, 101, 9045–9054, 1996.
- Solomon, P. M., Barrett, J., Connor, B. J., Zoonematkermani, S., Parrish, A., Lee, A., Pyle, J., and Chipperfield, M.: Seasonal observations of chlorine monoxide in the stratosphere over Antarctica during the 1996–1998 ozone holes and comparison with the SLIMCAT 3D model, *J. Geophys. Res.*, 105, 28979–29001, 2000.
- Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, X., Feist, D. G., and Wennberg, P. O.: The Total Carbon Column Observing Network's GGG2014 Data Version. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tcon.ggg2014.documentation.R0/1221662, 2015.